

Una sólida metodología complementaria que aporta sencillez: el árbol de decisión

por Édgar Zamora
FACTUM



Las herramientas metodológicas que se conforman a partir de dos ejes de variables han demostrado que, a pesar de su eficacia, son insuficientes para dar cuenta de la complejidad de un entorno analizado. A lo largo del presente artículo, Édgar Zamora nos muestra, a grandes rasgos, un modelo sencillo y claro, que nos permite inferir cuáles son los factores que actúan en el objeto de estudio y que, por su especificidad, bien puede aplicarse a los estudios de mercado.

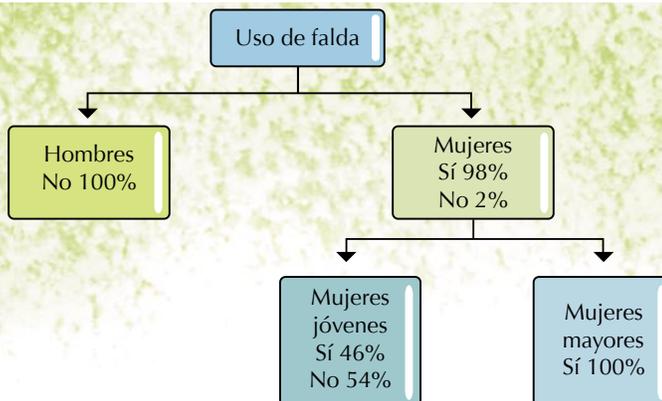
Para todos los que participamos de la vida moderna, resulta evidente que la sociedad se ha vuelto más compleja y dinámica, que los gustos, hábitos y prácticas de consumo se han transformado radicalmente en un plazo relativamente corto. Es también evidente que si tratamos de explicar. Las modificaciones a partir de una simple comparación de dos variables —las llamadas “técnicas vibariadas”—obtendremos un contexto insuficiente, limitado. Como respuesta a semejante problemática, han emergido métodos más avanzados, que nos ayudan a comprender mejor cómo se desenvuelven los clientes en el competido mundo en el que vivimos.

Dentro de esta línea de análisis presento una técnica sencilla de aplicar y que puede ser enormemente útil para obtener, en un tiempo relativamente corto, una idea muy completa de cómo segmentar los datos que obtenemos.

A qué llamamos árbol de decisión

Básicamente, se trata de un modelo representado de manera gráfica por un diagrama lógico. En éste, se divide una variable o pregunta de interés —llamada “dependiente”— de acuerdo con su relación con otras —también llamadas “predictoras” o “independientes” por su influencia en aquélla.

La estructura de dicho árbol se compone, además, de una base de datos que permite hacer ciertas predicciones. La característica fundamental de estas estimaciones consiste en que se basan en reglas, pruebas de medidas y pasos sucesivos. En resumen, su planteamiento se fundamenta en la experiencia. Se forman, así, “ramas” donde las variables independientes dividen a la dependiente, de tal forma que es posible observar la representación del problema, así como las categorías que se proponen para resolverlo.



Un ejemplo de la vida cotidiana

Para demostrar la utilidad del árbol de decisión, pongamos un ejemplo meramente descriptivo: en un pequeño pueblo se desea conocer quién lleva determinado tipo de ropa. Simplifiquemos seleccionando una prenda muy fácil de segmentar: la falda. Podríamos disponer de varias posibles variables para determinar quiénes usan falda; sin embargo, el mejor pronosticador es el sexo, pues prácticamente ningún hombre porta semejante atuendo.

La segmentación, de acuerdo con este ejemplo, implicaría no contentarse con una sola variable y definir otras que ayuden a distinguir con mayor claridad a las mujeres que usan faldas. De entrada —y considerando que es en una población pequeña donde efectuamos nuestro estudio—, la edad podría ser un criterio importante: es difícil ver a las mujeres mayores con pantalones, mientras que las jóvenes gustan mucho de usar este atuendo.

Lo anterior nos conduce a pensar que, de no introducirse nuevas variables, la población de nuestra comunidad de estudio quedaría segmentada en tres grupos: el de los hombres —en éste nadie usa faldas—; el de las mujeres mayores —el cual es muy probable que use faldas—, y el de las jóvenes —entre quienes existe un porcentaje medio de portadoras de faldas. Así, sin mayores dificultades y aplicando criterios muy claros, se trazan los grupos que permitirán, en una fase posterior, desarrollar los árboles.

Algunos datos adicionales de esta metodología

Los árboles de decisión se basan en diferentes técnicas, la más utilizada es la que se conoce como CHAID —Chi Square Automatic Interaction Detection—, misma que, como su nombre lo indica, emplea la prueba de χ^2 —chi cuadrada— para medir diferencias existentes. Por lo general, estos procedimientos se emplean con variables de tipo nominal —es decir, aquéllas cuyos valores de respuesta son categorías o descripciones.

Uno de los pasos más importantes —y también más difíciles— de la metodología del árbol de decisión consiste en determinar la variable de interés para el análisis —variable dependiente— y elegir un conjunto de posibles pronosticadores relevantes —variables independientes que repercuten en la anterior. Si se tiene alguna duda para agregar alguna variable independiente, lo que procede es, de todas maneras, incluirla, puesto que el mismo *software* se encargará de eliminarla de manera automática si no es idónea —lo que hace, por ejemplo, Answer Tree de SPSS.

Para resumir este punto: en comparación con otros análisis que controlan múltiples variables independientes de manera simultánea, el árbol examina una a la vez y divide segmentos de la dependiente, de tal manera que ésta no sea utilizada en lo sucesivo.

Parte de la coherencia y sencillez que aporta este modelo radica en su estructura, que establece un diagrama lógico en el cual una pregunta dependiente ayuda a desarrollar ramificaciones en las que actúan variables denominadas como “independientes” o “predictoras”.

El árbol de decisión, además de adaptable y lógico, ofrece la ventaja de ser sumamente comprensible para el estudioso. De ahí que puedan agregarse los niveles que sean necesarios y que demande el objeto de investigación.

Las fortalezas del árbol

El usuario puede comprobar la racionalidad del modelo y, si lo cree necesario, modificar el árbol o influir en su arquitectura. Claro que, cualquier cambio dependerá de la experiencia de quien lo aplique. Por ejemplo, el número de niveles del árbol puede ser decisión de quien vaya a trabajar con él, lo cual apunta a una absoluta flexibilidad.

Otra virtud del método radica en que permite la incorporación de valores perdidos, agrupándolos como si se tratara de otro valor común, atributo que refuerza la flexibilidad antes mencionada.

Sus puntos más delgados

CHAID requiere de grandes cantidades de datos para asegurarse de que las observaciones de los grupos son significativas.

Y sus aplicaciones

Los árboles de decisión pueden tener distintas aplicaciones; algunas de ellas son:

- ◆ Captar clientes potenciales, utilizando criterios como: la satisfacción actual con los proveedores, el nivel de facturación, las preferencias por determinadas líneas, etcétera.
- ◆ Distinguir entre los clientes que dejan ganancia y los que no.
- ◆ Ubicar a los clientes leales de los que no lo son.
- ◆ Comparar la respuesta a ciertos productos de clientes interesados y de quienes no mostraron entusiasmo alguno.

Soluciones como anillo al dedo



En CINCO hemos desarrollado una gama de productos y servicios puntuales de análisis y estrategia para potenciar el éxito de nuestros clientes:

CiNCO

CONSULTORES EN INVESTIGACIÓN
Y COMUNICACIÓN, S. A.

- Climas y tendencias electorales
- Geomarketing - Político
- Manejo de crisis
- Representaciones sociales
- Diagnóstico de servicio
- Constelación del clima organizacional
- Reputación corporativa
- Calidad receptiva
- Geomarketing - Comercial Urbanístico

Con Cinco estás en buenas manos

CiNCO

CONSULTORES EN INVESTIGACIÓN Y COMUNICACIÓN, S.C.

AUDIENCIAS MERCADOS ORGANIZACIONAL OPINIÓN PÚBLICA

Río Churubusco 422 Col. Del Carmen Coyoacán México, D.F. 04100

Tel/Fax (52 55) 5658 3705 5658 3888 5659 5163 www.e-cinco.com.mx